

1FWO

**CRF Errors Edited by the STIC Systems  
Branch**

Serial Number: 10/785,673

CRF Edit Date: 3/8/04  
Edited by: ME

**ENTERED**

\_\_\_ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

\_\_\_ Corrected the SEQ ID NO. Sequence numbers edited were:

\_\_\_\_\_

\_\_\_ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

\_\_\_\_\_

\_\_\_ Deleted: \_\_\_ invalid beginning/end-of-file text ; \_\_\_ page numbers

\_\_\_ Inserted mandatory headings/numeric identifiers, specifically:

\_\_\_\_\_

\_\_\_ Moved responses to same line as heading/numeric identifier, specifically:

\_\_\_\_\_

✓ Other: Sequence 4 - corrected amino acid numbering

\_\_\_\_\_

\_\_\_\_\_

Revised 09/09/2003



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/08/2004

TIME: 15:22:42

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\03082004\J785673.raw

## SEQUENCE LISTING

```

4 (1) GENERAL INFORMATION:
C--> 6 (i) APPLICANT: MINETTI, CONCEICAO;
7 MICHON, FRANCIS;
8 PULLEN, JEFFREY K.;
9 POLDVINO-BODNAR, MARYELLEN;
10 LIANG, SHU-MEI;
11 TAI, JOSEPH Y.
13 (ii) TITLE OF INVENTION: MODIFIED IMMUNOGENIC
14 PNEUMOLYSIN COMPOSITIONS AS VACCINES
16 (iii) NUMBER OF SEQUENCES: 18
18 (iv) CORRESPONDENCE ADDRESS:
19 (A) ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
20 (B) STREET: 345 PARK AVENUE
21 (C) CITY: NEW YORK
22 (D) STATE: NEW YORK
23 (E) COUNTRY: USA
24 (F) ZIP: 10154
26 (v) COMPUTER READABLE FORM:
27 (A) MEDIUM TYPE: FLOPPY DISK
28 (B) COMPUTER: IBM PC COMPATIBLE
29 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
30 (D) SOFTWARE: MICROSOFT WORD 97
32 (vi) CURRENT APPLICATION DATA:
C--> 33 (A) APPLICATION NUMBER: US/10/785,673
C--> 34 (B) FILING DATE: 23-Feb-2004
35 (C) CLASSIFICATION:
41 (vii) PRIOR APPLICATION DATA:
38 (A) APPLICATION NUMBER: 60/053,306
39 (B) FILING DATE: 1997-07-21
42 (A) APPLICATION NUMBER: 60/073,456
43 (B) FILING DATE: 1998-02-02
45 (viii) ATTORNEY/AGENT INFORMATION:
46 (A) NAME: DARRYL H. STEENSMA
47 (B) REGISTRATION NUMBER: 43,155
49 (C) REFERENCE/DOCKET NUMBER: 1758-4036US2
51 (ix) TELECOMMUNICATION INFORMATION:
52 (A) TELEPHONE: (212) 758-4800
53 (B) TELEFAX: (212) 751-6849
54 (C) TELEX: 421792
57 (2) INFORMATION FOR SEQ ID NO: 1:
59 (i) SEQUENCE CHARACTERISTICS:
60 (A) LENGTH: 1413

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/08/2004

TIME: 15:22:42

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\03082004\J785673.raw

61 (B) TYPE: nucleic acid  
 62 (C) STRANDEDNESS: double  
 63 (D) TOPOLOGY: linear  
 65 (ii) MOLECULE TYPE: cDNA  
 67 (vi) ORIGINAL SOURCE:  
 68 (A) ORGANISM: S. pneumoniae  
 70 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

C--> 72 ATGGCAAATA AAGCAGTAAA TGACTTTATA CTAGCTATGA 40  
 74 ATTACGATAA AAAGAAACTC TTGACCCATC AGGGAGAAAG 80  
 76 TATTGAAAAT CGTTTCATCA AAGAGGGTAA TCAGCTACCC 120  
 78 GATGAGTTTG TTGTTATCGA AAGAAAGAAG CGGAGCTTGT 160  
 80 CGACAAATAC AAGTGATATT TCTGTAACAG CTACCAACGA 200  
 82 CAGTCGCCTC TATCCTGGAG CACTTCTCGT AGTGGATGAG 240  
 84 ACCTTGTTAG AGAATAATCC CACTCTTCTT GCGGTCGATC 280  
 86 GTGCTCCGAT GACTTATAGT ATTGATTTGC CTGGTTTGGC 320  
 88 AAGTAGCGAT AGCTTTCTCC AAGTGGAAGA TCCCAGCAAT 360  
 90 TCAAGTGTTT GCGGAGCGGT AAACGATTTG TTGGCTAAGT 400  
 92 GGCATCAAGA TTATGGTCAG GTCAATAATG TCCCAGCTAG 440  
 94 AATGCAGTAT GAAAAAATCA CGGCTCACAG CATGGAACAA 480  
 97 CTCAAGGTCA AGTTTGGTTC TGACTTTGAA AAGACAGGGA 520  
 99 ATTCTCTTGA TATTGATTTT AACTCTGTCC ATTCAGGCCA 560  
 101 AAAGCAGATT CAGATTGTTA ATTTTAAGCA GATTTATTAT 600  
 103 ACAGTCAGCG TAGACGCTGT TAAAAATCCA GGAGATGTGT 640  
 105 TTCAAGATAC TGTAACGGTA GAGGATTTAA AACAGAGAGG 680  
 107 AATTTCTGCA GAGCGTCCTT TGGTCTATAT TTCGAGTGTT 720  
 109 GCTTATGGGC GCCAAGTCTA TCTCAAGTTG GAAACCACGA 760  
 111 GTAAGAGTGA TGAAGTAGAG GCTGCTTTTG AAGCTTTGAT 800  
 113 AAAAGGAGTC AAGGTAGCTC CTCAGACAGA GTGGAAGCAG 840  
 115 ATTTTGGACA ATACAGAAGT GAAGGCGGTT ATTTTAGGGG 880  
 117 GCGACCCAAG TTCGGGTGCC CGAGTTGTAA CAGGCAAGGT 920  
 119 GGATATGGTA GAGGACTTGA TTCAAGAAGG CAGTCGCTTT 960  
 121 ACAGCAGATC ATCCAGGCTT GCCGATTTCC TATACAACCT 1000  
 123 CTTTTTTACG TGACAATGTA GTTGCGACCT TTCAAATAG 1040  
 125 TACAGACTAT GTTGAGACTA AGGTTACAGC TTACAGAAAC 1080  
 127 GGAGATTTAC TGCTGGATCA TAGTGGTGCC TATGTTGCCC 1120  
 129 AATATTATAT TACTTGGAAT GAATTATCCT ATGATCATCA 1160  
 131 AGGTAAGGAA GTCTTGACTC CTAAGGCTTG GGACAGAAAT 1200  
 133 GGGCAGGATT TAACGGCTCA CTTTACCACT AGTATTCCTT 1240  
 135 TAAAAGGGAA TGTTTCGTAAT CTCTCTGTCA AAATTAGAGA 1280  
 137 GTGTACCGGG CTTGCTTGGG AATGGTGGCG TACGGTTTAT 1320  
 139 GAAAAAACCG ATTTGCCACT AGTGCGTAAG CGGACGATTT 1360  
 141 CTATTTGGGG AACAACCTC TATCCGCAGG TAGAAGATAA 1400  
 144 GGTAAGAAAT GAC 1413

147 (2) INFORMATION FOR SEQ ID NO: 2:  
 149 (i) SEQUENCE CHARACTERISTICS:  
 150 (A) LENGTH: 1413  
 151 (B) TYPE: nucleic acid  
 152 (C) STRANDEDNESS: double  
 153 (D) TOPOLOGY: linear

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/08/2004

TIME: 15:22:42

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\03082004\J785673.raw

155 (ii) MOLECULE TYPE: cDNA

157 (vi) ORIGINAL SOURCE:

158 (A) ORGANISM: S. pneumoniae

160 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

C--> 162 ATGGCAAATA AAGCAGTAAA TGACTTTATA CTAGCTATGA 40  
 164 ATTACGATAN AAANAACTC TTGACCCATC AGGGAGAAAG 80  
 166 TATTGAAAAT CGTTTCANCA AAGAGGGTAA TCAGCTACCC 120  
 168 GNTGAGTTTG TTGNTANCGA AAGAAAGAAG CGGAGCTTGT 160  
 170 CGACAAATAC AAGTGATATT NCTGTANCAG CTACCNACGA 200  
 172 CAGTCGCCTC TATCCTGGAG CACTTCTCGT AGTGGATGAG 240  
 174 ACCTTGTNAG AGAATAATCC CACTCTTCTT GCGGTNGATC 280  
 176 GTGCTCCGAT GACTTATAGT ANTGNNTTGC CTGGTTTGGC 320  
 178 AAGTAGCGAT AGCTTTCTCC AAGTGGAAGA NCCCAGCAAT 360  
 180 TCAAGTGTTT GCGGAGCGGN ANACGATTG TTGGCTAAGT 400  
 182 GGCATCAAGA TTATGGTCAG GTCAATAATG TCCAGCTAG 440  
 184 AANGCAGTAT GAAAAAATNA CGGCTCACAG CATGGAACAA 480  
 186 CTCAAGGTCA AGTTTGGTTC TGACTTTGAA AAGNCAGGGA 520  
 188 ATTCTCTTGA TATTGATTTT AACTCTGTCC ATTCAAGNGA 560  
 191 AAAGCNGATT CAGATTGTTA ATNTTAAGCA GATTTATTAT 600  
 193 ACAGTCAGCG TAGACGCTGT TAAAAATCCA GGAGATGTGT 640  
 195 TTCAAGATAC TGTAACGGTA GAGGATTAA AACAGAGAGG 680  
 197 AATTTCTGCA GAGCGTCCTT TGGTCTATAT TCGAGNGTT 720  
 199 GCTTATGGGC GCCAAGTCTA TCTCAAGTTG GAAACCACGA 760  
 201 GTANGAGTGN TGAAGTAGAG GCTGCTTTTG AAGCTTTGAT 800  
 203 AAAAGGAGTC AAGGTAGCTC CTCAGACAGA GTGGAAGCAG 840  
 205 ATTTTGGACA ATACAGAAGT GAAGGCGGTT ATTTTAGGGG 880  
 207 GCGACCCAAG TTCGGGTGCC CGAGTTGTAA CAGGCAAGGT 920  
 209 GGATATGGTA GAGGACTTGA TTCAAGAAGG CAGTCGCTTT 960  
 211 ACAGCAGATC ATCCAGGCTT GCCGATTTC TATACAACTT 1000  
 213 CTTTTTTACG TGACAATGTA GTTGCGACCT TTCAAAANAG 1040  
 215 TACAGACTAT GTTGAGACTA AGGTTACAGC TTACAGAAAC 1080  
 217 GGAGATTTAC TGCTGGATCA TAGTGGTGCC TATGTTGCCC 1120  
 219 AATATTATAT TACTTGGNAT GAATTATCCT ATGATCATCA 1160  
 221 AGGTAAGGAA GTCTTGACTC CTAAGGCTTG GGACAGAAAT 1200  
 223 GGCAGGATT TNACGGCTCA CTTTACCACT AGTATTCCTT 1240  
 225 TAAAAGGGAA TGTTGTAAT CTCTCTGTCA AAATTAGAGA 1280  
 227 GTGTACCGGG CTTGCNTGGG AATGGTGGCG TACGGTTTAT 1320  
 229 GAAAAAACCG ATTTGCCACT AGTGCGTAAG CGGACGATTT 1360  
 231 CTATTTGGGG AACAACCTC TATCCNCAGG TAGANGATAA 1400  
 233 GGTAGAAAAT GAC 1413

237 (2) INFORMATION FOR SEQ ID NO: 3:

239 (i) SEQUENCE CHARACTERISTICS:

240 (A) LENGTH: 471

241 (B) TYPE: amino acid

242 (C) STRANDEDNESS: unknown

243 (D) TOPOLOGY: linear

W--&gt; 245 (ii) MOLECULE TYPE: amino acid

247 (vi) ORIGINAL SOURCE:

248 (A) ORGANISM: S. pneumoniae

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/08/2004

TIME: 15:22:42

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\03082004\J785673.raw

```

250      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
252 Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala
253 1          5          10
254 Met Asn Tyr Asp Lys Lys Lys Leu Leu Thr His Gln
255          15          20
256 Gly Glu Ser Ile Glu Asn Arg Phe Ile Lys Glu Gly
257 25          30          35
258 Asn Gln Leu Pro Asp Glu Phe Val Val Ile Glu Arg
259          40          45
260 Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile
261          50          55          60
262 Ser Val Thr Ala Thr Asn Asp Ser Arg Leu Tyr Pro
263          65          70
264 Gly Ala Leu Leu Val Val Asp Glu Thr Leu Leu Glu
265          75          80
266 Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro
267 85          90          95
268 Met Thr Tyr Ser Ile Asp Leu Pro Gly Leu Ala Ser
269          100          105
270 Ser Asp Ser Phe Leu Gln Val Glu Asp Pro Ser Asn
271          110          115          120
272 Ser Ser Val Arg Gly Ala Val Asn Asp Leu Leu Ala
273          125          130
274 Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val
275          135          140
276 Pro Ala Arg Met Gln Tyr Glu Lys Ile Thr Ala His
277 145          150          155
278 Ser Met Glu Gln Leu Lys Val Lys Phe Gly Ser Asp
279          160          165
281 Phe Glu Lys Thr Gly Asn Ser Leu Asp Ile Asp Phe
282          170          175          180
283 Asn Ser Val His Ser Gly Glu Lys Gln Ile Gln Ile
284          185          190
285 Val Asn Phe Lys Gln Ile Tyr Tyr Thr Val Ser Val
286          195          200
287 Asp Ala Val Lys Asn Pro Gly Asp Val Phe Gln Asp
288 205          210          215
289 Thr Val Thr Val Glu Asp Leu Lys Gln Arg Gly Ile
290          220          225
291 Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Ser Val
292          230          235          240
293 Ala Tyr Gly Arg Gln Val Tyr Leu Lys Leu Glu Thr
294          245          250
295 Thr Ser Lys Ser Asp Glu Val Glu Ala Ala Phe Glu
296          255          260
297 Ala Leu Ile Lys Gly Val Lys Val Ala Pro Gln Thr
298 265          270          275
299 Glu Trp Lys Gln Ile Leu Asp Asn Thr Glu Val Lys
300          280          285

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/08/2004

TIME: 15:22:42

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\03082004\J785673.raw

```

301 Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala
302      290                      295                      300
303 Arg Val Val Thr Gly Lys Val Asp Met Val Glu Asp
304                      305                      310
305 Leu Ile Gln Glu Gly Ser Arg Phe Thr Ala Asp His
306      315                      320
307 Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
308 325                      330                      335
309 Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr
310      340                      345
311 Asp Tyr Val Glu Thr Lys Val Thr Ala Tyr Arg Asn
312 350                      355                      360
313 Gly Asp Leu Leu Leu Asp His Ser Gly Ala Tyr Val
314      365                      370
315 Ala Gln Tyr Tyr Ile Thr Trp Asn Glu Leu Ser Tyr
316      375                      380
317 Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala
318 385                      390                      395
319 Trp Asp Arg Asn Gly Gln Asp Leu Thr Ala His Phe
320      400                      405
321 Thr Thr Ser Ile Pro Leu Lys Gly Asn Val Arg Asn
322 410                      415                      420
323 Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
324      425                      430
325 Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp
326      435                      440
328 Leu Pro Leu Val Arg Lys Arg Thr Ile Ser Ile Trp
329 445                      450                      455
330 Gly Thr Thr Leu Tyr Pro Gln Val Glu Asp Lys Val
331      460                      465
332 Glu Asn Asp
333      470

```

336 (2) INFORMATION FOR SEQ ID NO: 4:

338 (i) SEQUENCE CHARACTERISTICS:

339 (A) LENGTH: 471

340 (B) TYPE: amino acid

341 (C) STRANDEDNESS: unknown

342 (D) TOPOLOGY: linear

W--&gt; 344 (ii) MOLECULE TYPE: amino acid

346 (vi) ORIGINAL SOURCE:

347 (A) ORGANISM: S. pneumoniae

349 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

351 Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala
352 1                      5                      10

```

```

W--> 353 Met Asn Tyr Asp Xaa Xaa Lys Leu Leu Thr His Gln
354      15                      20

```

```

355 Gly Glu Ser Ile Glu Asn Arg Phe Xaa Lys Glu Gly
356 25                      30                      35

```

```

357 Asn Gln Leu Pro Xaa Glu Phe Val Xaa Xaa Glu Arg

```



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/05/2004

TIME: 12:01:54

Input Set : A:\17584036.txt

Output Set: N:\CRF4\03052004\J785673.raw

## SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

C--> 6 (i) APPLICANT: MINETTI, CONCEICAO;  
7 MICHON, FRANCIS;  
8 PULLEN, JEFFREY K.;  
9 POLDVINO-BODNAR, MARYELLEN;  
10 LIANG, SHU-MEI;  
11 TAI, JOSEPH Y.

13 (ii) TITLE OF INVENTION: MODIFIED IMMUNOGENIC  
14 PNEUMOLYSIN COMPOSITIONS AS VACCINES

16 (iii) NUMBER OF SEQUENCES: 18

18 (iv) CORRESPONDENCE ADDRESS:  
19 (A) ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
20 (B) STREET: 345 PARK AVENUE  
21 (C) CITY: NEW YORK  
22 (D) STATE: NEW YORK  
23 (E) COUNTRY: USA  
24 (F) ZIP: 10154

26 (v) COMPUTER READABLE FORM:  
27 (A) MEDIUM TYPE: FLOPPY DISK  
28 (B) COMPUTER: IBM PC COMPATIBLE  
29 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
30 (D) SOFTWARE: MICROSOFT WORD 97

32 (vi) CURRENT APPLICATION DATA:  
C--> 33 (A) APPLICATION NUMBER: US/10/785,673  
C--> 34 (B) FILING DATE: 23-Feb-2004  
35 (C) CLASSIFICATION:

41 (vii) PRIOR APPLICATION DATA:  
38 (A) APPLICATION NUMBER: 60/053,306  
39 (B) FILING DATE: 1997-07-21  
42 (A) APPLICATION NUMBER: 60/073,456  
43 (B) FILING DATE: 1998-02-02

45 (viii) ATTORNEY/AGENT INFORMATION:  
46 (A) NAME: DARRYL H. STEENSMA  
47 (B) REGISTRATION NUMBER: 43,155  
49 (C) REFERENCE/DOCKET NUMBER: 1758-4036US2

51 (ix) TELECOMMUNICATION INFORMATION:  
52 (A) TELEPHONE: (212) 758-4800  
53 (B) TELEFAX: (212) 751-6849  
54 (C) TELEX: 421792

P.2

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/05/2004

TIME: 12:01:54

Input Set : A:\17584036.txt

Output Set: N:\CRF4\03052004\J785673.raw

## ERRORED SEQUENCES

336 (2) INFORMATION FOR SEQ ID NO: 4:  
 338 (i) SEQUENCE CHARACTERISTICS:  
 339 (A) LENGTH: 471  
 340 (B) TYPE: amino acid  
 341 (C) STRANDEDNESS: unknown  
 342 (D) TOPOLOGY: linear  
 W--> 344 (ii) MOLECULE TYPE: amino acid  
 346 (vi) ORIGINAL SOURCE:  
 347 (A) ORGANISM: S. pneumoniae  
 349 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 351 Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala  
 352 1 5 10  
 W--> 353 Met Asn Tyr Asp Xaa Xaa Lys Leu Leu Thr His Gln  
 354 15 20  
 355 Gly Glu Ser Ile Glu Asn Arg Phe Xaa Lys Glu Gly  
 356 25 30 35  
 357 Asn Gln Leu Pro Xaa Glu Phe Val Xaa Xaa Glu Arg  
 358 40 45  
 359 Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile  
 360 50 55 60  
 361 Xaa Val Xaa Ala Thr Xaa Asp Ser Arg Leu Tyr Pro  
 362 65 70  
 363 Gly Ala Leu Leu Val Val Asp Glu Thr Xaa Leu Glu  
 364 75 80  
 365 Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro  
 366 85 90 95  
 367 Met Thr Tyr Ser Xaa Xaa Leu Pro Gly Leu Ala Ser  
 368 100 105  
 369 Ser Asp Ser Phe Leu Gln Val Glu Asp Pro Ser Asn  
 370 110 115 120  
 371 Ser Ser Val Arg Gly Ala Xaa Xaa Asp Leu Leu Ala  
 372 125 130  
 373 Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val  
 374 135 140  
 376 Pro Ala Arg Xaa Gln Tyr Glu Lys Xaa Thr Ala His  
 E--> 377 145 150 155 155 155  
 378 Ser Met Glu Gln Leu Lys Val Lys Phe Gly Ser Asp  
 E--> 379 160 165  
 380 Phe Glu Lys Xaa Gly Asn Ser Leu Asp Ile Asp Phe  
 E--> 381 170 175 180  
 382 Asn Ser Val His Ser Gly Glu Lys Xaa Ile Gln Ile  
 E--> 383 185 190  
 384 Val Asn Xaa Lys Gln Ile Tyr Tyr Thr Val Ser Val  
 E--> 385 195 200  
 386 Asp Ala Val Lys Asn Pro Gly Asp Val Phe Gln Asp  
 E--> 387 205 210 215  
 388 Thr Val Thr Val Glu Asp Leu Lys Gln Arg Gly Ile  
 E--> 389 220 225



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/785,673

DATE: 03/05/2004

TIME: 12:01:54

Input Set : A:\17584036.txt

Output Set: N:\CRF4\03052004\J785673.raw

```

390 Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Xaa Val
E--> 391      230      235      240
392 Ala Tyr Xaa Arg Gln Val Tyr Leu Lys Leu Glu Thr
E--> 393      245      250
394 Thr Ser Xaa Ser Xaa Glu Val Glu Ala Ala Phe Glu
E--> 395      255      260
396 Ala Leu Ile Lys Gly Val Lys Val Ala Pro Gln Thr
E--> 397 265      270      275
398 Glu Trp Lys Gln Ile Leu Asp Asn Thr Xaa Val Lys
E--> 399      280      285
400 Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala
E--> 401      290      295      300
402 Arg Val Val Thr Gly Lys Val Asp Met Val Glu Asp
E--> 403      305      310
404 Leu Ile Gln Glu Gly Ser Arg Phe Thr Ala Asp His
E--> 405      315      320
406 Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
E--> 407 325      330      335
408 Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr
E--> 409      340      345
410 Asp Tyr Val Glu Thr Lys Val Thr Ala Tyr Arg Asn
E--> 411      350      355      360
412 Gly Asp Leu Leu Leu Asp His Ser Gly Ala Tyr Val
E--> 413      365      370
414 Ala Gln Tyr Tyr Ile Thr Trp Xaa Glu Leu Ser Tyr
E--> 415      375      380
416 Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala
E--> 417 385      390      395
418 Trp Asp Arg Asn Gly Gln Asp Leu Thr Ala His Phe
E--> 419      400      405
420 Thr Thr Ser Ile Pro Leu Lys Gly Asn Val Arg Asn
E--> 421      410      415      420
423 Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
E--> 424      425      430
425 Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp
E--> 426      435      440
427 Leu Xaa Leu Val Arg Lys Arg Thr Ile Ser Ile Trp
E--> 428 445      450      455
429 Gly Thr Thr Leu Tyr Pro Gln Val Glu Asp Lys Val
E--> 430      460      465
431 Glu Asn Asp
E--> 432      470

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/785,673

DATE: 03/05/2004

TIME: 12:01:55

Input Set : A:\17584036.txt

Output Set: N:\CRF4\03052004\J785673.raw

L:6 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]  
L:33 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:34 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:72 M:111 C: (47) String data converted to upper case,  
M:111 Repeated in SeqNo=1  
L:162 M:111 C: (47) String data converted to upper case,  
M:111 Repeated in SeqNo=2  
L:245 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3  
L:344 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4  
L:353 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:12  
M:341 Repeated in SeqNo=4  
L:377 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:4  
M:332 Repeated in SeqNo=4  
L:443 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5  
L:450 M:111 C: (47) String data converted to upper case,  
L:461 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6  
L:468 M:111 C: (47) String data converted to upper case,  
L:480 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7  
L:487 M:111 C: (47) String data converted to upper case,  
L:498 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8  
L:505 M:111 C: (47) String data converted to upper case,  
L:517 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9  
L:524 M:111 C: (47) String data converted to upper case,  
L:535 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10  
L:542 M:111 C: (47) String data converted to upper case,  
L:553 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11  
L:560 M:111 C: (47) String data converted to upper case,  
L:572 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12  
L:579 M:111 C: (47) String data converted to upper case,  
L:590 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=13  
L:597 M:111 C: (47) String data converted to upper case,  
L:609 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14  
L:616 M:111 C: (47) String data converted to upper case,  
L:626 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15  
L:633 M:111 C: (47) String data converted to upper case,  
L:644 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=16  
L:651 M:111 C: (47) String data converted to upper case,  
L:663 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=17  
L:670 M:111 C: (47) String data converted to upper case,  
L:681 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=18  
L:688 M:111 C: (47) String data converted to upper case,